Massive datasets of patient information can be extremely useful during a pandemic, allowing public health officials to track outbreaks of disease, helping clinical caregivers and researchers to understand how effective different treatments are, or identifying individuals who are immune and do not pose risks to individuals or the community. Compiling data from multiple sources poses significant technical challenges, of course. But in the fraught atmosphere of a pandemic, it’s important to remember that the creation and use of such datasets also pose challenges for balancing the interests of individual data subjects and the wider community, just as they do under normal circumstances.

The AMA Code of Medical Ethics sets out the bedrock obligation to protect the confidentiality of information “gathered in association with the care of the patient” in Opinion 3.2.1, “Confidentiality.” Inappropriate disclosure of personal health information may expose the patient to risks of discrimination, stigmatization or other harms. Opinions elsewhere in the Code provide guidance about conditions under which patient information may ethically be used for purposes other than clinical care.

Most relevant is Opinion 7.3.7, “Safeguards in the use of DNA databanks,” which provides that patients “who contribute to research involving DNA databanks have a right to be informed about the nature and scope of the research and to make decisions about how their information may be used.” Physicians/researchers must obtain appropriately informed consent, disclosing the purposes for which the patient’s information will be used, the privacy standards to which research will adhere and that they have right to withdraw at any time. Disclosure must also include any conflicts of interest the researcher has, and whether researchers or data subjects benefit financially.

Use of datasets for commercial purposes raises questions about who holds property rights in patient data, equitable distribution of profits. Opinion 7.3.8, “Commercial use of human biological materials,” addresses these issues in the specific context of research with stored biological specimens, but its guidance is more broadly applicable.
Using datasets to identify the populations most vulnerable to infection or poor outcomes during a pandemic can provide important information, for example, for targeting public health or clinical interventions. Yet such use also carries the risk of stigmatizing groups, or further stigmatizing already minoritized or otherwise marginalized communities, beyond the possible harms of the disease itself. Guidance in Opinion 7.3.3, “International research,” states that research not exploit vulnerable populations and strive instead to lay the foundation for lasting benefit for the community.

Similarly, during an infectious disease outbreak when datasets are mined to identify individuals who are immune, care should be taken to ensure that any benefits conferred by that status, such as release from mandatory quarantine or isolation, are distributed equitably.

For additional policy on clinical registries from the AMA House of Delegates see H-450.933, "Clinical data registries."

**Additional ethics guidance in a pandemic**

The AMA offers an overview of foundational guidance regarding medical ethics for health care professionals and institutions responding to the COVID-19 pandemic.